Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Journal of Computational and Applied Mathematics 223 (2009) 646–658
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

An improved penalty method for power-law Stokes problems

Jeff Borggaard[a], Traian Iliescu[a], John Paul Roop[b,*]

[a] *Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0123, USA*
[b] *Department of Mathematics, North Carolina A&T State University, Greensboro, NC 27411, USA*

Received 28 September 2006; received in revised form 16 December 2007

Abstract

For the numerical approximation of fluid flow phenomena, it is often highly desirable to decouple the equations defining conservation of momentum and conservation of mass by using a penalty function method. The current penalty function methods for power-law Stokes fluids converge at a sublinear rate with respect to the penalty parameter. In this article, we show theoretically and numerically that a *linear* penalty function approximation to a power-law Stokes problem yields a higher-order accuracy over the known nonlinear penalty method. Theoretically, finite element approximation of the linear penalty function method is shown to satisfy an *improved* order of approximation with respect to the penalty parameter. The numerical experiments presented in the paper support the theoretical results and satisfy a *linear* order of approximation.

Keywords: Penalty method; Power-law Stokes; Numerical analysis; Fluids; Large eddy simulation; Smagorinsky model

1. Introduction

In this article we analyze mathematically and numerically an alternative penalty method for the solution of stationary *power-law Stokes problem*

$$\begin{cases} -\nu\, \nabla \cdot (|\nabla \boldsymbol{u}|^{r-2} \nabla \boldsymbol{u}) + \nabla p = \mathbf{f}, & \text{in } \Omega, \\ \nabla \cdot \boldsymbol{u} = 0, & \text{in } \Omega, \end{cases} \tag{1}$$

where $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ is the computational domain, ν the kinematic viscosity, $\boldsymbol{u}$ the velocity vector, p the pressure, $r > 1$, and $\mathbf{f}$ the body force. The power-law Stokes equations (1) have been used as a mathematical model in numerous applications of non-Newtonian flows: in chemical engineering [8], design of extrusion of dies [20], the study of lithosphere [12], and other geophysical applications [27]. Equally important, the power-law Stokes equations (1) can be thought of as a simplified, linearized setting for the *Smagorinsky model* [26], one of the most popular models in the *large eddy simulation* of turbulent flows [7,24].

* Corresponding author. Tel.: +1 336 334 7822, fax: +1 336 256 0876.
E-mail addresses: jborggaard@vt.edu (J. Borggaard), iliescu@vt.edu (T. Iliescu), jproop@ncat.edu (J.P. Roop).
URLs: http://www.math.vt.edu/people/borggajt (J. Borggaard), http://www.math.vt.edu/people/iliescu (T. Iliescu), http://www.ncat.edu/~jproop (J.P. Roop).

0377-0427/$ - see front matter
doi:10.1016/j.cam.2008.02.002

The *penalty method* decouples the calculation of velocity and pressure in the numerical discretization of Stokes and Navier–Stokes equations. The penalty method has been used extensively in the numerical simulation of fluid flows [11,15,22] because of its computational efficiency. The mathematical and numerical analyses for the penalty method applied to the Stokes and Navier–Stokes equations have been also provided [28,16,25].

Lefton and Wei considered in a series of papers [17,29,18] the following *nonlinear penalty method*:

$$\begin{cases} -\nu\,\nabla\cdot(|\nabla \boldsymbol{u}^\varepsilon|^{r-2}\nabla\boldsymbol{u}^\varepsilon)+\nabla p^\varepsilon=\mathbf{f}, & \text{in } \Omega,\\ (|\nabla\cdot\boldsymbol{u}^\varepsilon|^{r-2})\nabla\cdot\boldsymbol{u}^\varepsilon+\varepsilon\, p^\varepsilon=0, & \text{in } \Omega, \end{cases} \tag{2}$$

where ε is a small penalty parameter (typically $\varepsilon \sim 10^{-4}$–10^{-3} in practical calculations). The authors have proved error estimates of the form

$$\|\boldsymbol{u}-\boldsymbol{u}^\varepsilon\|_{1,r} \le \begin{cases} C\,\varepsilon^{\frac{1}{(r-1)^2}}, & \text{if } r\ge 2,\\ C\,\varepsilon^{\frac{1}{(3-r)(r-1)}}, & \text{if } 1<r\le 2, \end{cases} \tag{3}$$

where C is a generic constant independent of ε. Similar estimates were also derived for the stationary power-law Navier–Stokes equations [29]. Furthermore, the authors have considered a finite element discretization of the nonlinear penalty method applied to power-law Stokes equations (2) and proved error estimates of the form

$$\|\boldsymbol{u}-\boldsymbol{u}_h^\varepsilon\|_{1,r} \le \begin{cases} C\left[\left(\frac{1}{\varepsilon}\right)^{\frac{1}{r(r-1)}}\|\boldsymbol{u}-\boldsymbol{v}_h\|_{1,r}^{\frac{1}{r-1}}+\|\boldsymbol{u}-\boldsymbol{v}_h\|_{1,r}^{\frac{1}{r}}+\varepsilon^{\frac{1}{r^2}}\right], & \text{if } r\ge 2,\\ C\left[\left(\frac{1}{\varepsilon}\right)^{\frac{1}{r(3-r)}}\|\boldsymbol{u}-\boldsymbol{v}_h\|_{1,r}^{\frac{1}{3-r}}+\|\boldsymbol{u}-\boldsymbol{v}_h\|_{1,r}^{\frac{1}{2}}+\varepsilon^{\frac{1}{2r}}\right], & \text{if } 1<r\le 2, \end{cases} \tag{4}$$

for all $\boldsymbol{v}_h$ in the finite element space, where C is a generic constant independent of ε and h.

In this paper, we consider the following *linear penalty method* for the power-law Stokes equations (1)

$$\begin{cases} -\nu\,\nabla\cdot(|\nabla \boldsymbol{u}^\varepsilon|^{r-2}\nabla\boldsymbol{u}^\varepsilon)+\nabla p^\varepsilon=\mathbf{f}, & \text{in } \Omega,\\ \nabla\cdot\boldsymbol{u}^\varepsilon+\varepsilon\, p^\varepsilon=0, & \text{in } \Omega. \end{cases} \tag{5}$$

We only focus on the $r \ge 2$ case, although some of our results carry over to the $1 < r \le 2$ case. The solution $\boldsymbol{u}^\varepsilon$ to the continuous linear penalty method (5) is shown to satisfy the following error estimate

$$\|\boldsymbol{u}-\boldsymbol{u}^\varepsilon\|_{1,r} \le C\,\varepsilon^{1/r}, \tag{6}$$

where C is a generic constant independent of ε. Note that estimate (6) represents a significant improvement over the corresponding error estimate (3) satisfied by the nonlinear penalty method (2). Furthermore, we show that the solution $\boldsymbol{u}^\varepsilon$ to the finite element discretization of (5) satisfies the following *improved* error estimate

$$\|\boldsymbol{u}_h-\boldsymbol{u}_h^\varepsilon\|_{1,r} \le C\,h^{\frac{(2-r)d}{r(r-1)}}\,\varepsilon^{\frac{1}{r-1}}, \tag{7}$$

where $\boldsymbol{u}_h, \boldsymbol{u}_h^\varepsilon$ are finite element approximations of the continuous solutions $\boldsymbol{u}, \boldsymbol{u}^\varepsilon$ to (1) and (5), respectively, and C is a generic constant independent of ε and h. This is a significant improvement over the corresponding error estimate (4) satisfied by the finite element discretization of the nonlinear penalty method (2). We also note that a similar linear penalty method has been successfully used in [23] to approximate power-law flows.

The paper is organized as follows: In Section 2, we introduce the notation and functional spaces used in the paper. In Section 3, we introduce the improved linear penalty method for the power-law Stokes flows (5) and prove the existence and uniqueness of solutions. In Section 4, we present improved error estimates with respect to the penalty parameter ε for the continuous linear penalty method (5). In Section 5, we prove existence and uniqueness of solutions as well as improved error estimates with respect to ε for the finite element discretization of the linear penalty method (5). In Section 6, we present numerical experiments for the power-law Stokes problem (5), which illustrate the improved rate of convergence for the linear penalty method over the nonlinear penalty method. Finally, in Section 7 we present conclusions and directions of future research.

2. Notation and mathematical setting

We present below some of the notations and functional analysis results that will be frequently used in the paper.

Let $L^r(\Omega)$, $W^{k,r}(\Omega)$, and $W_0^{k,r}(\Omega)$, $1 < r < \infty, k = 0, 1, 2, \dots$ denote the usual Sobolev spaces [1]. Let $\|\cdot\|$ denote the norm on $L^2(\Omega)$, $\|\cdot\|_r = \|\cdot\|_{0,r}$ the norm on $L^r(\Omega)$, and $\|\cdot\|_{k,r}$ the norm on $W^{k,r}(\Omega)$. Let $(\cdot,\cdot)$ denote the scalar product in $L^2(\Omega)$. The vector spaces and vector functions will be indicated by boldface type letters. For $1 < r < \infty$, let r' denote the conjugate of r: $\left(\frac{1}{r} + \frac{1}{r'} = 1, \text{ i.e. } \quad r' = \frac{r}{r-1}\right)$. Let $W^{-k,r'}(\Omega)$ denote the dual space of $W_0^{k,r}(\Omega)$, and $\|\cdot\|_{-k,r'}$ the norm on $W^{-k,r'}(\Omega)$. Let $L_0^r(\Omega) := \left\{q \in L^r(\Omega) : \int_\Omega q \, \mathrm{d}\mathbf{x} = 0\right\}$.

We will also use the strong monotonicity and Lipschitz continuity of the r-Laplacian [19,21]:

$$\begin{aligned}&\left(|\nabla \boldsymbol{u}_1|^{r-2}\,\nabla \boldsymbol{u}_1, \nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)\right) - \left(|\nabla \boldsymbol{u}_2|^{r-2}\,\nabla \boldsymbol{u}_2, \nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)\right)\\ &\quad \geq C\,\|\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)\|_{0,r}^r \quad \forall \boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbf{W}^{1,r}(\Omega),\end{aligned} \tag{8}$$

$$\left(|\nabla \boldsymbol{u}_1|^{r-2}\,\nabla \boldsymbol{u}_1, \nabla \boldsymbol{v}\right) - \left(|\nabla \boldsymbol{u}_2|^{r-2}\,\nabla \boldsymbol{u}_2, \nabla \boldsymbol{v}\right) \leq C\,M\,\|\nabla(\boldsymbol{u}_1 - \boldsymbol{u}_2)\|_{0,r}\,\|\nabla \boldsymbol{v}\|_{0,r} \quad \forall \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{v} \in \mathbf{W}^{1,r}(\Omega), \tag{9}$$

where $M = \max\left\{\|\nabla \boldsymbol{u}_1\|_{0,r}^{r-2}, \|\nabla \boldsymbol{u}_2\|_{0,r}^{r-2}\right\}$, C is a generic constant depending on d, r, and Ω, but not on $\boldsymbol{u}_1, \boldsymbol{u}_2$ or $\boldsymbol{v}$. In the following, we will also assume $r \geq 2$, although some of our results carry over to the $r < 2$ case.

In addition, we will use the following interpolation result [9]:

Lemma 2.1. *Let $\{T_h\}$ $0 < h \leq 1$ denote a quasi-uniform family of subdivisions of a polyhedral domain $\Omega \subset \mathbb{R}^d$. Let $(\hat{K}, P, N)$ be a reference finite element such that $P \subset W^{l,p}(\hat{K}) \cap W^{m,q}(\hat{K})$ where $1 \leq p, q \leq \infty$ and $0 \leq m \leq l$. For $K \in T_h$, let (K, P_K, N_K) be the affine equivalent element, and $V_h = \{v : v$ is measurable and $v|_K \in P_K, \forall K \in T_h\}$. Then there exists $C = C(l, p, q)$ such that*

$$\left[\sum_{K \in T_h} \|v\|_{W^{l,p}(K)}^p\right]^{1/p} \leq C h^{m-l+\min\left(0, \frac{d}{p} - \frac{d}{q}\right)} \left[\sum_{K \in T_h} \|v\|_{W^{m,q}(K)}^q\right]^{1/q}. \tag{10}$$

3. The linear penalty method

In this section, we will introduce the improved, linear penalty method for the stationary power-law Stokes problem (1). We will also prove that there exists a unique solution to the linear penalty method.

Let $\mathbf{X} := \mathbf{W}_0^{1,r}(\Omega)$ and $Q := L_0^{r'}(\Omega)$.

The *mixed weak formulation* of the stationary power-law Stokes problem (1) reads

$$\begin{cases}\nu\,(|\nabla \boldsymbol{u}|^{r-2}\,\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) - (p, \nabla\cdot\boldsymbol{v}) = (\mathbf{f}, \boldsymbol{v}), & \forall\,\boldsymbol{v} \in \mathbf{X},\\ (\nabla\cdot\boldsymbol{u}, q) = 0, & \forall\,q \in Q.\end{cases} \tag{11}$$

The existence and uniqueness of solutions $(\boldsymbol{u}, p) \in \mathbf{X} \times Q$ to (11) were studied in [4,6].

The *mixed weak formulation* of the linear penalty method applied to the stationary power-law Stokes problem (5) reads

$$\begin{cases}\nu\,(|\nabla \boldsymbol{u}^\varepsilon|^{r-2}\,\nabla \boldsymbol{u}^\varepsilon, \nabla \boldsymbol{v}) - (p^\varepsilon, \nabla\cdot\boldsymbol{v}) = (\mathbf{f}, \boldsymbol{v}), & \forall\,\boldsymbol{v} \in \mathbf{X},\\ (\nabla\cdot\boldsymbol{u}^\varepsilon, q) + \varepsilon(p^\varepsilon, q) = 0, & \forall\,q \in Q.\end{cases} \tag{12}$$

To study the existence and uniqueness of solutions $(\boldsymbol{u}^\varepsilon, p^\varepsilon) \in \mathbf{X} \times Q$ to (12), we follow [17,6,10].

First, define the functional $J_\varepsilon : \mathbf{X} \to \mathbb{R}$ given by

$$J_\varepsilon(\boldsymbol{u}) := \frac{\nu}{r}\,\|\nabla \boldsymbol{u}\|_{0,r}^r + \frac{1}{2\,\varepsilon}\,\|\nabla\cdot\boldsymbol{u}\|^2 - (\mathbf{f}, \boldsymbol{u}), \quad \forall \boldsymbol{u} \in \mathbf{X}. \tag{13}$$

Note that J_ε is well defined. Indeed, since $\boldsymbol{u} \in \mathbf{X} = \mathbf{W}_0^{1,r}(\Omega)$ and $r \geq 2$, it follows that $\nabla\cdot\boldsymbol{u} \in L^2(\Omega)$.

Let $\langle \cdot, \cdot \rangle$ denote the duality pairing between $\mathbf{X}$ and $\mathbf{X}^*$. It is a straightforward calculation to check that $J_\varepsilon(\cdot)$ is Gâteaux differentiable on $\mathbf{X}$. Indeed,

$$\langle J'_\varepsilon(\boldsymbol{u}), \boldsymbol{v} \rangle = \frac{\mathrm{d}J_\varepsilon}{\mathrm{d}t}(\boldsymbol{u} + t\,\boldsymbol{v}) = \langle A\boldsymbol{u}, \boldsymbol{v} \rangle - \langle \mathbf{f}, \boldsymbol{v} \rangle, \quad \forall\, \boldsymbol{u}, \boldsymbol{v} \in \mathbf{X}, \tag{14}$$

where $A : \mathbf{X} \to \mathbf{X}^*$ is such that

$$\langle A\boldsymbol{u}, \boldsymbol{v} \rangle := \nu\,(|\nabla \boldsymbol{u}|^{r-2}\,\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + \frac{1}{\varepsilon}\,(\nabla \cdot \boldsymbol{u}, \nabla \cdot \boldsymbol{v}), \quad \forall\, \boldsymbol{u}, \boldsymbol{v} \in \mathbf{X}, \tag{15}$$

Moreover, $J'_\varepsilon(\cdot)$ is strictly monotone. Indeed,

$$\begin{aligned}
\langle J'_\varepsilon(\boldsymbol{w}_1) - J'_\varepsilon(\boldsymbol{w}_2), \boldsymbol{w}_1 - \boldsymbol{w}_2 \rangle &= \langle A\boldsymbol{w}_1 - A\boldsymbol{w}_2, \boldsymbol{w}_1 - \boldsymbol{w}_2 \rangle \\
&= \nu\,(|\nabla \boldsymbol{w}_1|^{r-2}\,\nabla \boldsymbol{w}_1 - |\nabla \boldsymbol{w}_2|^{r-2}\,\nabla \boldsymbol{w}_2, \nabla(\boldsymbol{w}_1 - \boldsymbol{w}_2)) \\
&\quad + \frac{1}{\varepsilon}\,(\nabla \cdot (\boldsymbol{w}_1 - \boldsymbol{w}_2), \nabla \cdot (\boldsymbol{w}_1 - \boldsymbol{w}_2)) \qquad (16)\\
&\geq C\,\|\boldsymbol{w}_1 - \boldsymbol{w}_2\|_{1,r}^r, \qquad (17)
\end{aligned}$$

where in the last inequality we used the strong monotonicity of the r-Laplacian (8). Therefore, $J_\varepsilon(\cdot)$ is strictly convex. (See Section 3 in [4].) Furthermore, $J_\varepsilon(\cdot)$ is coercive on $\mathbf{X}$. Thus, we have proved the following result:

Lemma 3.1. *There exists a unique solution to the minimization problem*

$$\min_{\boldsymbol{v} \in \mathbf{X}} J_\varepsilon(\boldsymbol{v}). \tag{18}$$

Remark 3.1. The minimization problem (18) is the *variational formulation* of (5), the linear penalty method for the power-law Stokes problem.

We will prove now that the existence and uniqueness of the solution to the variational formulation (18) imply the existence and uniqueness of solutions to the mixed weak formulation of the power-law Stokes problem (12). First, we note that the Euler–Lagrange equation of (18) is

$$\langle A\boldsymbol{u}, \boldsymbol{v} \rangle = (\mathbf{f}, \boldsymbol{v}), \quad \forall\, \boldsymbol{v} \in \mathbf{X}. \tag{19}$$

Let $(\boldsymbol{u}^\varepsilon, p^\varepsilon)$ be a solution of (12). Then, by using (12_2) with

$$q := \frac{1}{|\Omega|} \int_\Omega \nabla \cdot \boldsymbol{v}\, \mathrm{d}\mathbf{x} - \nabla \cdot \boldsymbol{v} \in L^r \subset L^{r'} \quad (\text{since } r \geq 2), \tag{20}$$

we obtain

$$\langle A\boldsymbol{u}^\varepsilon, \boldsymbol{v} \rangle = (\mathbf{f}, \boldsymbol{v}), \quad \forall\, \boldsymbol{v} \in \mathbf{X}, \tag{21}$$

which implies that $\boldsymbol{u}^\varepsilon$ is a solution to (19). Conversely, let $\boldsymbol{u}$ be a solution to (19). Pick $\boldsymbol{u}^\varepsilon := \boldsymbol{u}$ and

$$p^\varepsilon := \frac{1}{\varepsilon\,|\Omega|} \int_\Omega \nabla \cdot \boldsymbol{u}^\varepsilon\, \mathrm{d}\mathbf{x} - \frac{1}{\varepsilon}\,(\nabla \cdot \boldsymbol{u}^\varepsilon). \tag{22}$$

It is a simple calculation to check that $(\boldsymbol{u}^\varepsilon, p^\varepsilon)$ is a solution to (12).

In order to prove the uniqueness of p^ε, however, we require the Ladyzhenskaya–Babuska–Brezzi (LBB) inf–sup condition [14,15]. Amrouche and Girault [2] proved the following result:

Theorem 3.1. *Let Ω be a bounded, connected, Lipschitz continuous domain in $\mathbb{R}^d$ and let r be any real number with $1 < r < \infty$, and r' its conjugate. There exists a constant $\beta > 0$ such that*

$$0 < \beta \leq \inf_{q \in L_0^{r'}(\Omega)} \sup_{\boldsymbol{v} \in \mathbf{W}_0^{1,r}(\Omega)} \frac{(\nabla \cdot \boldsymbol{v}, q)}{\|q\|_{0,r'}\, \|\boldsymbol{v}\|_{1,r}}. \tag{23}$$

In the sequel, we will assume that the domain Ω satisfies the conditions in Theorem 3.1, and thus the LBB condition is satisfied. We are now in the position to prove the uniqueness of p^ε. Assume that $p_1^\varepsilon, p_2^\varepsilon \in Q$ satisfy (12). We know that $\boldsymbol{u}^\varepsilon$ is unique (Lemma 3.1). Thus, (12_2) for p_1^ε and p_2^ε yields

$$\varepsilon\,(p_1^\varepsilon - p_2^\varepsilon, \nabla\cdot\boldsymbol{v}) = 0 \quad \forall\, \boldsymbol{v} \in \mathbf{W}_0^{1,r}. \tag{24}$$

By picking $q := p_1^\varepsilon - p_2^\varepsilon$ in (23), we get

$$\sup_{\boldsymbol{v}\in\mathbf{W}_0^{1,r}(\Omega)} \frac{(\nabla\cdot\boldsymbol{v}, p_1^\varepsilon - p_2^\varepsilon)}{\|\boldsymbol{v}\|_{1,r}} \geq \beta\,\|p_1^\varepsilon - p_2^\varepsilon\|_{0,r'}. \tag{25}$$

Thus, (24) and (25) imply $0 \geq \beta\|p_1^\varepsilon - p_2^\varepsilon\|_{0,r'}$, leading to the uniqueness of p^ε.

Thus, since $\varepsilon > 0$, we have proved the following result:

Theorem 3.2. *There exists a unique solution $(\boldsymbol{u}^\varepsilon, p^\varepsilon)$ to the mixed weak formulation of the linear penalty method for the power-law Stokes problem (12). Moreover, the unique pressure $p^\varepsilon \in L_0^r(\Omega)$ is given by the following formula*

$$p^\varepsilon = \frac{1}{\varepsilon\,|\Omega|}\int_\Omega \nabla\cdot\boldsymbol{u}^\varepsilon\,\mathrm{d}\mathbf{x} - \frac{1}{\varepsilon}\,(\nabla\cdot\boldsymbol{u}^\varepsilon). \tag{26}$$

4. Error analysis for the continuous linear penalty method

In this section, we prove that $(\boldsymbol{u}^\varepsilon, p^\varepsilon)$, the solution to the continuous linear penalty method (12) converges to $(\boldsymbol{u}, p)$, the solution to the power-law Stokes problem (11) strongly with respect to ε.

We start by proving an *a priori* bound for $\boldsymbol{u}$ and $\boldsymbol{u}^\varepsilon$.

Lemma 4.1. *Let $(\boldsymbol{u}, p)$ be the solution of (11) and $(\boldsymbol{u}^\varepsilon, p^\varepsilon)$ the solution of (12). Then $\|\boldsymbol{u}\|_{1,r} \leq C$ and $\|\boldsymbol{u}^\varepsilon\|_{1,r} \leq C$, where C is a generic constant depending on r, Ω, ν, and $\mathbf{f}$, but not on ε.*

Proof. The *a priori* bound for $\boldsymbol{u}$ was proved in [17]. Setting $\boldsymbol{v} := \boldsymbol{u}^\varepsilon$ in (12), and using the strong monotonicity of the r-Laplacian (8) and the Cauchy–Schwartz inequality, prove the lemma. □

The next theorem proves the strong convergence of $\boldsymbol{u}^\varepsilon$ to $\boldsymbol{u}$ as $\varepsilon \to 0$ in the $\mathbf{W}^{1,r}(\Omega)$ norm, provided that p is contained in $L^2(\Omega)$.

Theorem 4.1. *Let $(\boldsymbol{u}, p)$ solve (11) and $(\boldsymbol{u}^\varepsilon, p^\varepsilon)$ solve (12). Then, provided that $p \in L^2(\Omega)$, $\boldsymbol{u}^\varepsilon \to \boldsymbol{u}$ as $\varepsilon \to 0$ strongly in the $\mathbf{W}^{1,r}(\Omega)$ norm. Specifically,*

$$\|\boldsymbol{u} - \boldsymbol{u}^\varepsilon\|_{1,r} \leq C\,\varepsilon^{1/r}. \tag{27}$$

where C is a generic constant depending on ν, but not on ε.

Proof. First, by subtracting (12_1) from (11_1), we get

$$\nu\,(|\nabla\boldsymbol{u}|^{r-2}\nabla\boldsymbol{u}, \nabla\boldsymbol{v}) - \nu\,(|\nabla\boldsymbol{u}^\varepsilon|^{r-2}\nabla\boldsymbol{u}^\varepsilon, \nabla\boldsymbol{v}) - (p - p^\varepsilon, \nabla\cdot\boldsymbol{v}) = 0. \tag{28}$$

By setting $\boldsymbol{v} := \boldsymbol{u} - \boldsymbol{u}^\varepsilon$ in (28), we obtain

$$\nu\,(|\nabla\boldsymbol{u}|^{r-2}\nabla\boldsymbol{u}, \nabla(\boldsymbol{u}-\boldsymbol{u}^\varepsilon)) - \nu\,(|\nabla\boldsymbol{u}^\varepsilon|^{r-2}\nabla\boldsymbol{u}^\varepsilon, \nabla(\boldsymbol{u}-\boldsymbol{u}^\varepsilon)) - (p - p^\varepsilon, \nabla\cdot(\boldsymbol{u}-\boldsymbol{u}^\varepsilon)) = 0. \tag{29}$$

Next, by using the continuity equation (11_2), (29) becomes

$$\nu\,(|\nabla\boldsymbol{u}|^{r-2}\nabla\boldsymbol{u}, \nabla(\boldsymbol{u}-\boldsymbol{u}^\varepsilon)) - \nu\,(|\nabla\boldsymbol{u}^\varepsilon|^{r-2}\nabla\boldsymbol{u}^\varepsilon, \nabla(\boldsymbol{u}-\boldsymbol{u}^\varepsilon)) + (p - p^\varepsilon, \nabla\cdot\boldsymbol{u}^\varepsilon) = 0. \tag{30}$$

By using (12_2), Eq. (30) reads

$$\nu\,(|\nabla\boldsymbol{u}|^{r-2}\nabla\boldsymbol{u}, \nabla(\boldsymbol{u}-\boldsymbol{u}^\varepsilon)) - \nu\,(|\nabla\boldsymbol{u}^\varepsilon|^{r-2}\nabla\boldsymbol{u}^\varepsilon, \nabla(\boldsymbol{u}-\boldsymbol{u}^\varepsilon)) + \varepsilon(p^\varepsilon, p^\varepsilon) = \varepsilon(p, p^\varepsilon). \tag{31}$$

By Theorem 3.2, $p^\varepsilon \in L^r(\Omega)$. Since $r \geq 2$, this implies that $p^\varepsilon \in L^2(\Omega)$. Next, by using the strong monotonicity of the r-Laplacian (8), the regularity of p^ε, the regularity assumption $p \in L^2(\Omega)$, and Young's inequality, we have

$$\nu \, \|\boldsymbol{u} - \boldsymbol{u}^\varepsilon\|_{1,r}^r + \varepsilon \, \|p^\varepsilon\|^2 \leq \frac{\varepsilon}{2} \, \|p\|^2 + \frac{\varepsilon}{2} \, \|p^\varepsilon\|^2, \tag{32}$$

from which (27) clearly follows. □

5. Error analysis for the finite element approximation of the linear penalty method

In this section, we turn our attention to the finite element discretization of the linear penalty method (12). First, the existence and uniqueness result as well as the preliminary error bound follow, provided that the approximating subspaces satisfy the discrete LBB condition. Next, we prove that the discrete form of the linear penalty method possesses a *linear* order of convergence via the discrete Sobolev inequalities.

Let $\mathbf{X}_h \subset \mathbf{X} = \mathbf{W}_0^{1,r}(\Omega)$ and $Q_h \subset Q = L_0^{r'}(\Omega)$ be two conforming finite element spaces satisfying the *discrete LBB condition* [13–15]: There exists a constant $\widetilde{\beta} > 0$ *independent of* h such that

$$0 < \widetilde{\beta} \leq \inf_{q_h \in Q_h} \sup_{\boldsymbol{v} \in \mathbf{X}_h} \frac{(\nabla \cdot \boldsymbol{v}_h, q_h)}{\|q_h\|_{0,r'} \, \|\boldsymbol{v}_h\|_{1,r}}. \tag{33}$$

Remark 5.1. In all the numerical experiments in Section 6 we will use either the Taylor–Hood finite element pair or the mini-element. Note that both finite element pairs satisfy the discrete LBB condition (33) with $\widetilde{\beta}$ independent of h (Lemma 4.20 and Lemma 4.21 in [13], respectively).

The *mixed weak formulation* for the finite element discretization of the power-law Stokes problem (11) reads

$$\begin{cases} \nu \, (|\nabla \boldsymbol{u}_h|^{r-2} \, \nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) - (p_h, \nabla \cdot \boldsymbol{v}_h) = (\mathbf{f}, \boldsymbol{v}_h), & \forall \, \boldsymbol{v}_h \in \mathbf{X}_h, \\ (\nabla \cdot \boldsymbol{u}_h, q_h) = 0, & \forall \, q_h \in Q_h, \end{cases} \tag{34}$$

The existence and uniqueness of solutions to (34) was studied in [3,4,6,5].

The *mixed weak formulation* for the finite element discretization of the linear penalty method (12) reads:

Find $(\boldsymbol{u}_h^\varepsilon, p_h^\varepsilon) \in \mathbf{X}_h \times Q_h$ such that

$$\begin{cases} \nu \, (|\nabla \boldsymbol{u}_h^\varepsilon|^{r-2} \, \nabla \boldsymbol{u}_h^\varepsilon, \nabla \boldsymbol{v}_h) - (p_h^\varepsilon, \nabla \cdot \boldsymbol{v}_h) = (\mathbf{f}, \boldsymbol{v}_h), & \forall \, \boldsymbol{v}_h \in \mathbf{X}_h, \\ (\nabla \cdot \boldsymbol{u}_h^\varepsilon, q_h) + \varepsilon (p_h^\varepsilon, q_h) = 0, & \forall \, q_h \in Q_h. \end{cases} \tag{35}$$

Theorem 5.1. *Let $\mathbf{X}_h \subset \mathbf{X} = \mathbf{W}_0^{1,r}(\Omega)$ and $Q_h \subset Q = L_0^{r'}(\Omega)$ be two conforming finite element spaces satisfying the discrete LBB condition* (33). *Then, there exists a unique solution $(\boldsymbol{u}_h^\varepsilon, p_h^\varepsilon) \in \mathbf{X}_h \times Q_h$ to* (35).

Proof. The proof follows along the same lines as the proof of Theorem 3.2.

We start by constructing the functional $J_\varepsilon^h : \mathbf{X}_h \to \mathbb{R}$ given by

$$J_\varepsilon^h(\boldsymbol{u}_h) := \frac{\nu}{r} \, \|\nabla \boldsymbol{u}_h\|_{0,r}^r + \frac{1}{2\,\varepsilon} \, \|\nabla \cdot \boldsymbol{u}_h\|^2 - (\mathbf{f}, \boldsymbol{u}_h), \quad \forall \boldsymbol{u}_h \in \mathbf{X}_h. \tag{36}$$

Next, we prove that J_ε^h is strictly convex, and therefore the minimization problem

$$\min_{\boldsymbol{v}_h \in \mathbf{X}_h} J_\varepsilon^h(\boldsymbol{v}_h). \tag{37}$$

has a unique solution $\boldsymbol{u}_h^\varepsilon \in \mathbf{X}_h$. Then, by using the discrete LBB condition (33), we prove that (37) and (35) are equivalent. Since (37) has a unique solution, we infer that (35) has a unique solution as well, which proves the theorem. □

We now establish the main result, which is that (35), the discretization of the linear penalty method, satisfies the error estimate $\|\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon\|_{1,r} \leq C(h)\varepsilon^{1/(r-1)}$. We start by proving some supporting lemmas.

The first lemma consists of *a priori* bounds for $\boldsymbol{u}_h$, $\boldsymbol{u}_h^\varepsilon$, and p_h.

Lemma 5.1. *Let* $(\boldsymbol{u}_h, p_h)$ *solve* (34) *and* $(\boldsymbol{u}_h^\varepsilon, p_h^\varepsilon)$ *solve* (35). *Then* $\|\boldsymbol{u}_h\|_{1,r} \le C$, $\|\boldsymbol{u}_h^\varepsilon\|_{1,r} \le C$, *and* $\|p_h\|_{0,r'} \le C$, *where* C *is a generic constant depending on* r, Ω, ν, $\mathbf{f}$, *and* $\widetilde{\beta}$, *but not on* ε *and* h.

Proof. Setting $\boldsymbol{v}_h := \boldsymbol{u}_h$ in (34), and using the strong monotonicity of the r-Laplacian (8) and the Cauchy–Schwartz inequality, prove the *a priori* bound for $\boldsymbol{u}_h$.

Setting $\boldsymbol{v}_h := \boldsymbol{u}_h^\varepsilon$ in (35), and using the strong monotonicity of the r-Laplacian (8), and the Cauchy–Schwartz inequality, prove the *a priori* bound for $\boldsymbol{u}_h^\varepsilon$.

Finally, setting $q_h := p_h$ in the discrete LBB condition (33), using (34), the strong monotonicity of the r-Laplacian (8), and the Cauchy–Schwartz inequality, prove the *a priori* bound for p_h. □

The next lemma is an *a priori* bound for $\nabla \cdot \boldsymbol{u}_h^\varepsilon$.

Lemma 5.2. *Let* $(\boldsymbol{u}_h, p_h)$ *solve* (34) *and* $(\boldsymbol{u}_h^\varepsilon, p_h^\varepsilon)$ *solve* (35). *Then*

$$\|\nabla \cdot \boldsymbol{u}_h^\varepsilon\|_r \le C h^{\frac{(2-r)d}{r}} \varepsilon, \tag{38}$$

where C *is a generic constant depending on* r, Ω, *and* $\mathbf{f}$, *but not on* ε *and* h.

Proof. First, by subtracting (35_1) from (34_1), we get

$$\nu\,(|\nabla \boldsymbol{u}_h|^{r-2}\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) - \nu\,(|\nabla \boldsymbol{u}_h^\varepsilon|^{r-2}\nabla \boldsymbol{u}_h^\varepsilon, \nabla \boldsymbol{v}_h) - (p_h - p_h^\varepsilon, \nabla \cdot \boldsymbol{v}_h) = 0. \tag{39}$$

By setting $\boldsymbol{v}_h := \boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon$ in (39), we obtain

$$\nu\,(|\nabla \boldsymbol{u}_h|^{r-2}\nabla \boldsymbol{u}_h, \nabla(\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon)) - \nu\,(|\nabla \boldsymbol{u}_h^\varepsilon|^{r-2}\nabla \boldsymbol{u}_h^\varepsilon, \nabla(\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon)) - (p_h - p_h^\varepsilon, \nabla \cdot (\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon)) = 0. \tag{40}$$

Next, by using the continuity equation (34_2), (40) becomes

$$\nu\,(|\nabla \boldsymbol{u}_h|^{r-2}\nabla \boldsymbol{u}_h, \nabla(\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon)) - \nu\,(|\nabla \boldsymbol{u}_h^\varepsilon|^{r-2}\nabla \boldsymbol{u}_h^\varepsilon, \nabla(\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon)) + (p_h - p_h^\varepsilon, \nabla \cdot \boldsymbol{u}_h^\varepsilon) = 0. \tag{41}$$

From (35_2) and (41), we immediately have

$$\nu\,(|\nabla \boldsymbol{u}_h|^{r-2}\nabla \boldsymbol{u}_h, \nabla(\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon)) - \nu\,(|\nabla \boldsymbol{u}_h^\varepsilon|^{r-2}\nabla \boldsymbol{u}_h^\varepsilon, \nabla(\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon)) + \left(p_h + \frac{1}{\varepsilon}\nabla \cdot \boldsymbol{u}_h^\varepsilon, \nabla \cdot \boldsymbol{u}_h^\varepsilon\right) = 0. \tag{42}$$

Using the strong monotonicity of the r-Laplacian (8), we obtain

$$\left(p_h + \frac{1}{\varepsilon}\nabla \cdot \boldsymbol{u}_h^\varepsilon, \nabla \cdot \boldsymbol{u}_h^\varepsilon\right) \le 0. \tag{43}$$

Since $\boldsymbol{u}_h^\varepsilon \in \mathbf{W}^{1,r}(\Omega)$, we have $\nabla \cdot \boldsymbol{u}_h^\varepsilon \in L^r(\Omega)$. Thus, (43) and the Cauchy–Schwartz inequality yield

$$\frac{1}{\varepsilon}\,\|\nabla \cdot \boldsymbol{u}_h^\varepsilon\|_2^2 \le \|p_h\|_{r'}\,\|\nabla \cdot \boldsymbol{u}_h^\varepsilon\|_r. \tag{44}$$

Now, from the discrete Sobolev inequality, Lemma 2.1, we have

$$\|\nabla \cdot \boldsymbol{u}_h^\varepsilon\|_r \le C h^{\frac{(2-r)d}{2r}}\,\|\nabla \cdot \boldsymbol{u}_h^\varepsilon\|_2. \tag{45}$$

Thus, (44) and (45), and the *a priori* bound for p_h in Lemma 5.1 prove (38). □

Remark 5.2. Notice that Lemma 5.2 does not require that $p_h \in L^2(\Omega)$, as in Theorem 4.1. The only requirement is that $p_h \in Q = L_0^{r'}(\Omega)$.

The next lemma bounds the $\mathbf{W}^{0,r'}$ norm of $(p_h - p_h^\varepsilon)$ in terms of the $\mathbf{W}^{1,r}$ norm of $(\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon)$.

Lemma 5.3. *Let* $(\boldsymbol{u}_h, p_h)$ *solve* (34) *and* $(\boldsymbol{u}_h^\varepsilon, p_h^\varepsilon)$ *solve* (35). *Then*

$$\|p_h - p_h^\varepsilon\|_{0,r'} \le C\,\|\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon\|_{1,r}, \tag{46}$$

where C *is a generic constant depending on* ν, r, $\widetilde{\beta}$, Ω, *and* $\mathbf{f}$, *but not on* ε *and* h.

Proof. By letting $q := p_h - p_h^{\varepsilon} \in L_0^{r'}(\Omega)$ in the discrete LBB condition (33), we get

$$\|p_h - p_h^{\varepsilon}\|_{r'} \leq \widetilde{\beta}^{-1} \sup_{\boldsymbol{v}_h \in \mathbf{X}_h} \frac{(p_h - p_h^{\varepsilon}, \nabla \cdot \boldsymbol{v}_h)}{\|\boldsymbol{v}_h\|_{1,r}}. \tag{47}$$

By using (39), Eq. (47) becomes

$$\|p_h - p_h^{\varepsilon}\|_{r'} \leq \widetilde{\beta}^{-1} \sup_{\boldsymbol{v}_h \in \mathbf{X}_h} \frac{\nu\,(|\nabla \boldsymbol{u}_h|^{r-2}\nabla \boldsymbol{u}_h, \nabla \boldsymbol{v}_h) - \nu\,(|\nabla \boldsymbol{u}_h^{\varepsilon}|^{r-2}\nabla \boldsymbol{u}_h^{\varepsilon}, \nabla \boldsymbol{v}_h)}{\|\boldsymbol{v}_h\|_{1,r}}. \tag{48}$$

Finally, (46) follows from (48) and the Lipschitz continuity of the r-Laplacian (9). □

The next lemma bounds the $\mathbf{W}^{1,r}$ norm of $(\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon})$ in terms of the L^r norm of $\nabla \cdot \boldsymbol{u}_h^{\varepsilon}$.

Lemma 5.4. *Let $(\boldsymbol{u}_h, p_h)$ solve (34) and $(\boldsymbol{u}_h^{\varepsilon}, p_h^{\varepsilon})$ solve (35). Then*

$$\|\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon}\|_{1,r} \leq C\,\|\nabla \cdot \boldsymbol{u}_h^{\varepsilon}\|_r^{1/(r-1)}, \tag{49}$$

where C is a generic constant depending on ν, r, $\widetilde{\beta}$, Ω, and $\mathbf{f}$, but not on ε and h.

Proof. Eq. (41) implies that

$$\nu\,(|\nabla \boldsymbol{u}_h|^{r-2}\nabla \boldsymbol{u}_h, \nabla(\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon})) - \nu\,(|\nabla \boldsymbol{u}_h^{\varepsilon}|^{r-2}\nabla \boldsymbol{u}_h^{\varepsilon}, \nabla(\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon})) \leq \|p_h - p_h^{\varepsilon}\|_{r'}\,\|\nabla \cdot \boldsymbol{u}_h^{\varepsilon}\|_r. \tag{50}$$

By using Lemma 5.3, Eq. (50) becomes

$$\nu\,(|\nabla \boldsymbol{u}_h|^{r-2}\nabla \boldsymbol{u}_h, \nabla(\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon})) - \nu\,(|\nabla \boldsymbol{u}_h^{\varepsilon}|^{r-2}\nabla \boldsymbol{u}_h^{\varepsilon}, \nabla(\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon})) \leq C\,\|\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon}\|_{1,r}\,\|\nabla \cdot \boldsymbol{u}_h^{\varepsilon}\|_r. \tag{51}$$

Using the strong monotonicity of the r-Laplacian (8), Eq. (51) implies

$$\|\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon}\|_{1,r}^r \leq C\,\|\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon}\|_{1,r}\,\|\nabla \cdot \boldsymbol{u}_h^{\varepsilon}\|_r, \tag{52}$$

which proves (49). □

The next theorem, which is the main result of this paper, proves error estimates for the convergence of $(\boldsymbol{u}_h^{\varepsilon}, p_h^{\varepsilon})$ to $(\boldsymbol{u}_h, p_h)$.

Theorem 5.2. *Let Ω be a bounded, connected, Lipschitz continuous domain in $\mathbb{R}^d$. Let $2 \leq r < \infty$ and r' be its conjugate. Let $(\boldsymbol{u}_h, p_h)$ be the unique solution of (34), $\boldsymbol{u}^{\varepsilon}$ the unique solution of (37), and*

$$p_h^{\varepsilon} := \frac{1}{\varepsilon\,|\Omega|}\int_{\Omega} \nabla \cdot \boldsymbol{u}_h^{\varepsilon}\,\mathrm{d}\mathbf{x} - \frac{1}{\varepsilon}\,\nabla \cdot \boldsymbol{u}_h^{\varepsilon}. \tag{53}$$

Then, there exists a generic constant C, depending on ν, r, Ω, $\widetilde{\beta}$ and $\mathbf{f}$, but not on ε and h, such that

$$\|\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon}\|_{1,r} \leq C\,h^{\frac{(2-r)d}{r(r-1)}}\,\varepsilon^{\frac{1}{r-1}}, \quad \text{and} \tag{54}$$

$$\|p_h - p_h^{\varepsilon}\|_{0,r'} \leq C\,h^{\frac{(2-r)d}{r(r-1)}}\,\varepsilon^{\frac{1}{r-1}}. \tag{55}$$

Proof. The estimate (54) clearly follows from Lemmas 5.2 and 5.4. Next, the estimate (55) clearly follows from Lemma 5.3 and (54). □

6. Numerical experiments

In this section, we present numerical experiments that investigate the rate of convergence of $(\boldsymbol{u}_h^{\varepsilon}, p_h^{\varepsilon})$ to $(\boldsymbol{u}_h, p_h)$ *with respect to the penalty parameter* ε. Here $(\boldsymbol{u}_h, p_h)$ and $(\boldsymbol{u}_h^{\varepsilon}, p_h^{\varepsilon})$ are the finite element approximations of the power-law Stokes problem (34) and linear penalty method applied to the power-law Stokes problem (35), respectively. In our numerical experiments, we will investigate the rate of convergence of $(\boldsymbol{u}_h^{\varepsilon}, p_h^{\varepsilon})$ to $(\boldsymbol{u}_h, p_h)$, and not to $(\boldsymbol{u}, p)$. In order to investigate the convergence results presented in Theorem 5.2, we present convergence rates for $\|\boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon}\|_{1,r}$ and $\|p_h - p_h^{\varepsilon}\|_{0,r'}$.

Table 1
Velocity convergence results for the nonlinear penalty method (60) in Experiment 1, $r > 2$; grid parameter $h = 1/32$

ε	$\|\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon\|_{1,r}$ ($r = 3$)	Convergence rate	$\|\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon\|_{1,r}$ ($r = 7/2$)	Convergence rate	$\|\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon\|_{1,r}$ ($r = 4$)	Convergence rate
$10^{-3}/4$	1.006455×10^{-2}		2.831250×10^{-2}		5.577870×10^{-2}	
$10^{-3}/8$	6.822246×10^{-3}	0.56	2.136560×10^{-2}	0.41	4.446320×10^{-2}	0.33
$10^{-3}/16$	4.455791×10^{-3}	0.61	1.604230×10^{-2}	0.41	3.535730×10^{-2}	0.33
$10^{-3}/32$	2.780892×10^{-3}	0.68	1.194170×10^{-2}	0.43	2.804088×10^{-2}	0.33
$10^{-3}/64$	1.666585×10^{-3}	0.74	8.761141×10^{-3}	0.45	2.217126×10^{-2}	0.34
Predicted		0.25		0.16		0.11

As model problem, we considered the power-law Stokes problem with homogeneous Dirichlet boundary conditions

$$-\nu\nabla \cdot \left[(|\nabla \boldsymbol{u}|^{r-2})\nabla \boldsymbol{u}\right] + \nabla p = \mathbf{f}, \quad \text{in } \Omega \tag{56}$$

$$\nabla \cdot \boldsymbol{u} = 0, \quad \text{in } \Omega \tag{57}$$

$$\boldsymbol{u} = g, \quad \text{on } \partial\Omega, \tag{58}$$

where the computational domain $\Omega := (0, 1) \times (0, 1) \subset \mathbb{R}^2$. We utilize the right-hand side functions $\mathbf{f}$ corresponding to the true solution:

$$\boldsymbol{u} := \begin{bmatrix} -\cos(\pi x)\sin(\pi y) \\ \sin(\pi x)\cos(\pi y) \end{bmatrix}, \qquad p := -\frac{1}{4}\left(\cos(2\pi x) + \cos(2\pi y)\right). \tag{59}$$

We note that in order to investigate the rates of convergence for the pressure, (53) must be enforced computationally to ensure that $\int_\Omega p_h^\varepsilon \mathrm{d}x = 0$.

We present five computational experiments that indicate experimental convergence rates for both the nonlinear and linear penalty function method. Note that in all experiments performed, the linear penalty function method exhibited a linear convergence rate with respect to the penalty parameter ε.

Experiment 1. For this numerical experiment, we computed convergence rates for the *nonlinear penalty function method*, in which (57) is replaced by the nonlinear relationship

$$(|\nabla \cdot \boldsymbol{u}^\varepsilon|^{r-2})\nabla \cdot \boldsymbol{u}^\varepsilon + \varepsilon\, p^\varepsilon = 0. \tag{60}$$

We discretized the resulting system (56) and (60) using the Taylor–Hood finite element pair and used a uniform mesh with characteristic length $h = 1/32$. For the nonlinearities in (56) and (60), we employed the Newton iteration scheme

$$\begin{aligned}
&\nu\left(|\nabla \boldsymbol{u}_h^{\varepsilon(n-1)}|^{r-2}\nabla \boldsymbol{u}_h^{\varepsilon(n)}, \nabla \boldsymbol{v}_h\right) + \nu(r-2)\left(|\nabla \boldsymbol{u}_h^{\varepsilon(n-1)}|^{r-4}\left[\nabla \boldsymbol{u}_h^{\varepsilon(n-1)} : \nabla \boldsymbol{u}_h^{\varepsilon(n)}\right]\nabla \boldsymbol{u}_h^{\varepsilon(n-1)}, \nabla \boldsymbol{v}_h\right) \\
&\quad - \left(p_h^{\varepsilon(n)}, \nabla \cdot \boldsymbol{v}_h\right) = (\boldsymbol{f}, \boldsymbol{v}_h) + \nu(r-2)\left(|\nabla \boldsymbol{u}_h^{\varepsilon(n-1)}|^{r-2}\nabla \boldsymbol{u}_h^{\varepsilon(n-1)}, \nabla \boldsymbol{v}_h\right), \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{X}_h, \\
&(r-1)\left(|\nabla \cdot \boldsymbol{u}_h^{\varepsilon(n-1)}|^{r-2}\nabla \cdot \boldsymbol{u}_h^{\varepsilon(n)}, q_h\right) + \varepsilon\left(p_h^{\varepsilon(n)}, q_h\right) \\
&\quad = (r-2)\left(|\nabla \cdot \boldsymbol{u}_h^{\varepsilon(n-1)}|^{r-2}\nabla \cdot \boldsymbol{u}_h^{\varepsilon(n-1)}, q_h\right), \quad \forall\, q_h \in Q_h.
\end{aligned}$$

We performed computations for the model problem and $\nu = 10^{-2}$, $r = 3, 7/2, 4$. The results are summarized in Tables 1 and 2. Notice that the convergence rates for the velocity and pressure in the nonlinear penalty function observed in our computational experiments were on the order of $\frac{1}{(r-1)}$. In the sequel, we observe that the experimental convergence rates for the linear penalty function method are superior to the experimental convergence rates for the nonlinear penalty function method.

Experiment 2. For this numerical experiment, we computed convergence rates for the *linear penalty function method*, in which (57) is replaced by the linear relationship

$$\nabla \cdot \boldsymbol{u}^\varepsilon + \varepsilon\, p^\varepsilon = 0. \tag{61}$$

Table 2
Pressure convergence results for the nonlinear penalty method (60) in Experiment 1, $r > 2$; grid parameter $h = 1/32$

ε	$\Vert p_h - p_h^\varepsilon \Vert_{0,r'}$ ($r = 3$)	Convergence rate	$\Vert p_h - p_h^\varepsilon \Vert_{0,r'}$ ($r = 7/2$)	Convergence rate	$\Vert p_h - p_h^\varepsilon \Vert_{0,r'}$ ($r = 4$)	Convergence rate
$10^{-3}/4$	5.893092×10^{-4}		3.012872×10^{-3}		1.080856×10^{-2}	
$10^{-3}/8$	4.009367×10^{-4}	0.56	2.277192×10^{-3}	0.40	8.631992×10^{-3}	0.32
$10^{-3}/16$	2.622084×10^{-4}	0.61	1.712802×10^{-3}	0.41	6.874817×10^{-3}	0.33
$10^{-3}/32$	1.631244×10^{-4}	0.68	1.277546×10^{-3}	0.42	5.458722×10^{-3}	0.33
$10^{-3}/64$	9.656317×10^{-5}	0.76	9.390265×10^{-4}	0.44	4.318782×10^{-3}	0.34
Predicted		0.25		0.16		0.11

Table 3
Velocity convergence results for the linear penalty method (61) in Experiment 2, $r > 2$; grid parameter $h = 1/32$

ε	$\Vert \boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon \Vert_{1,r}$ ($r = 3$)	Convergence rate	$\Vert \boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon \Vert_{1,r}$ ($r = 7/2$)	Convergence rate	$\Vert \boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon \Vert_{1,r}$ ($r = 4$)	Convergence rate
$10^{-3}/4$	9.959125×10^{-5}		1.057540×10^{-4}		1.117104×10^{-4}	
$10^{-3}/8$	4.979656×10^{-5}	1.00	5.287920×10^{-5}	1.00	5.586032×10^{-5}	1.00
$10^{-3}/16$	2.489851×10^{-5}	1.00	2.644016×10^{-5}	1.00	2.793145×10^{-5}	1.00
$10^{-3}/32$	1.244931×10^{-5}	1.00	1.322022×10^{-5}	1.00	1.396605×10^{-5}	1.00
$10^{-3}/64$	6.224671×10^{-6}	1.00	6.610143×10^{-6}	1.00	6.983104×10^{-6}	1.00
Predicted		0.5		0.4		0.33

Table 4
Pressure convergence results for the linear penalty method (61) in Experiment 2, $r > 2$; grid parameter $h = 1/32$

ε	$\Vert p_h - p_h^\varepsilon \Vert_{0,r'}$ ($r = 3$)	Convergence rate	$\Vert p_h - p_h^\varepsilon \Vert_{0,r'}$ ($r = 7/2$)	Convergence rate	$\Vert p_h - p_h^\varepsilon \Vert_{0,r'}$ ($r = 4$)	Convergence rate
$10^{-3}/4$	6.036943×10^{-6}		1.215962×10^{-5}		2.452045×10^{-5}	
$10^{-3}/8$	3.018535×10^{-6}	1.00	6.080098×10^{-6}	1.00	1.226153×10^{-5}	1.00
$10^{-3}/16$	1.509284×10^{-6}	1.00	3.040121×10^{-6}	1.00	6.131093×10^{-6}	1.00
$10^{-3}/32$	7.546458×10^{-7}	1.00	1.520079×10^{-6}	1.00	3.065628×10^{-6}	1.00
$10^{-3}/64$	3.773239×10^{-7}	1.00	7.600439×10^{-7}	1.00	1.532835×10^{-6}	1.00
Predicted		0.5		0.4		0.33

We discretized the resulting system (56) and (61) using the Taylor–Hood finite element pair and used a uniform mesh with characteristic length $h = 1/32$. For the nonlinearity in (56), we employed the Newton iteration scheme

$$\nu \left(|\nabla \boldsymbol{u}_h^{\varepsilon(n-1)}|^{r-2} \nabla \boldsymbol{u}_h^{\varepsilon(n)}, \nabla \boldsymbol{v}_h \right) + \nu(r-2) \left(|\nabla \boldsymbol{u}_h^{\varepsilon(n-1)}|^{r-4} \left[\nabla \boldsymbol{u}_h^{\varepsilon(n-1)} : \nabla \boldsymbol{u}_h^{\varepsilon(n)} \right] \nabla \boldsymbol{u}_h^{\varepsilon(n-1)}, \nabla \boldsymbol{v}_h \right)$$
$$- \left(p_h^{\varepsilon(n)}, \nabla \cdot \boldsymbol{v}_h \right) = (\boldsymbol{f}, \boldsymbol{v}_h) + \nu(r-2) \left(|\nabla \boldsymbol{u}_h^{\varepsilon(n-1)}|^{r-2} \nabla \boldsymbol{u}_h^{\varepsilon(n-1)}, \nabla \boldsymbol{v}_h \right), \quad \forall\, \boldsymbol{v}_h \in \boldsymbol{X}_h, \tag{62}$$

$$\left(\nabla \cdot \boldsymbol{u}_h^{\varepsilon(n)}, q_h \right) + \varepsilon \left(p_h^{\varepsilon(n)}, q_h \right) = 0, \quad \forall\, q_h \in Q_h. \tag{63}$$

We performed computations for the model problem and $\nu = 10^{-2}$, $r = 3,\ 7/2,\ 4$. The results are summarized in Tables 3 and 4. Notice that the predicted convergence rates in the linear penalty function method for the velocity and pressure are both $\frac{1}{(r-1)}$. However, for our computational experiments, we observed velocity and pressure convergence rates of 1.

Experiment 3. For this numerical experiment, we investigated the h dependence in the theoretically predicted convergence rates in Theorem 5.2. For this experiment, we utilized the Taylor–Hood finite element pair, the Newton iteration scheme (63), and values $r = 3$, $\nu = 10^{-2}$, and $h = 1/16, 1/24, 1/32$, and $1/48$. The results of this computational experiment are presented in Table 5. We note that there is no particular dependency upon h in the observed convergence rates of 1.

Table 5
Velocity and pressure convergence results for the linear penalty method (61) in Experiment 3, $r = 3$; for various values of ε and grid parameters h

h	ε	$\|\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon\|_{1,r}$ ($r = 3$)	ε convergence rate	$\|p_h - p_h^\varepsilon\|_{0,r'}$ ($r = 3$)	ε convergence rate
1/16	$10^{-3}/4$	9.959190×10^{-5}		6.034270×10^{-6}	
1/16	$10^{-3}/8$	4.979688×10^{-5}	1.00	3.017199×10^{-6}	1.00
1/16	$10^{-3}/16$	2.489867×10^{-5}	1.00	1.508616×10^{-6}	1.00
1/16	$10^{-3}/32$	1.244939×10^{-5}	1.00	7.543116×10^{-7}	1.00
1/16	$10^{-3}/64$	6.224712×10^{-6}	1.00	3.771568×10^{-7}	1.00
1/24	$10^{-3}/4$	9.959376×10^{-5}		6.036514×10^{-6}	
1/24	$10^{-3}/8$	4.979781×10^{-5}	1.00	3.018321×10^{-6}	1.00
1/24	$10^{-3}/16$	2.489914×10^{-5}	1.00	1.509176×10^{-6}	1.00
1/24	$10^{-3}/32$	1.244963×10^{-5}	1.00	7.545922×10^{-7}	1.00
1/24	$10^{-3}/64$	6.224829×10^{-6}	1.00	3.772971×10^{-7}	1.00
1/32	$10^{-3}/4$	9.959125×10^{-5}		6.036943×10^{-6}	
1/32	$10^{-3}/8$	4.979656×10^{-5}	1.00	3.018535×10^{-6}	1.00
1/32	$10^{-3}/16$	2.489851×10^{-5}	1.00	1.509284×10^{-6}	1.00
1/32	$10^{-3}/32$	1.244931×10^{-5}	1.00	7.546458×10^{-7}	1.00
1/32	$10^{-3}/64$	6.224671×10^{-6}	1.00	3.773239×10^{-7}	1.00
1/48	$10^{-3}/4$	9.958831×10^{-5}		6.037192×10^{-6}	
1/48	$10^{-3}/8$	4.979509×10^{-5}	1.00	3.018660×10^{-6}	1.00
1/48	$10^{-3}/16$	2.489778×10^{-5}	1.00	1.509346×10^{-6}	1.00
1/48	$10^{-3}/32$	1.244895×10^{-5}	1.00	7.546770×10^{-7}	1.00
1/48	$10^{-3}/64$	6.224488×10^{-6}	1.00	3.773395×10^{-7}	1.00

Table 6
Velocity convergence results for the linear penalty method (61) in Experiment 4, $r = 3$; mini-element discretization; grid parameter $h = 1/32$

ε	$\|\boldsymbol{u}_h - \boldsymbol{u}_h^\varepsilon\|_{1,r}$ ($r = 3$)	Convergence rate	$\|p_h - p_h^\varepsilon\|_{0,r'}$ ($r = 3$)	Convergence rate
$10^{-3}/4$	1.001808×10^{-4}		6.180896×10^{-6}	
$10^{-3}/8$	5.009139×10^{-5}	1.00	3.090517×10^{-6}	1.00
$10^{-3}/16$	2.504594×10^{-5}	1.00	1.545276×10^{-6}	1.00
$10^{-3}/32$	1.252303×10^{-5}	1.00	7.726423×10^{-7}	1.00
$10^{-3}/64$	6.261530×10^{-6}	1.00	3.863222×10^{-7}	1.00

Experiment 4. For this numerical experiment, we computed convergence rates for the linear penalty function method for two other velocity–pressure finite element discretizations. We discretized (56) and (61) utilizing the mini-element instead of the Taylor–Hood finite element, which consists of a piecewise linear basis with a bubble function for the velocities and a piecewise linear basis for the pressure. For this experiment, we utilized the Newton iteration scheme (63) and values $\nu = 10^{-2}$, $r = 3$ $h = 1/32$. The results of this computational experiment are presented in Table 6. Notice that for this discretization, we also observed velocity and pressure convergence rates of 1.

Experiment 5. For this numerical experiment, we computed convergence rates for the linear penalty function method for values $r < 2$, in particular $r = 3/2$ and $7/4$. For this experiment, we utilized the Taylor–Hood finite element pair, the Newton iteration scheme (63), and values $r = 3/2, 7/4$, $\nu = 10^{-2}$, and $h = 1/32$. The results of this computational experiment are presented in Tables 7 and 8. Although values of r less than 2 are outside of the scope of our theoretical results, it is interesting to note that we observed linear experimental convergence rates for the velocity and pressure for this case as well.

7. Conclusions

In this paper, we have analyzed both mathematically and numerically, an improvement in the penalty method for the power-law Stokes problem. In particular, we proved that the finite element discretization of a *linear* penalty

Table 7
Velocity convergence results for the linear penalty method (61) in Experiment 5, $r < 2$; grid parameter $h = 1/32$

ε	$\Vert \boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon} \Vert_{1,r}$ $(r = 3/2)$	Convergence rate	$\Vert \boldsymbol{u}_h - \boldsymbol{u}_h^{\varepsilon} \Vert_{1,r}$ $(r = 7/4)$	Convergence rate
$10^{-3}/4$	8.595249×10^{-5}		8.721107×10^{-5}	
$10^{-3}/8$	4.294630×10^{-5}	1.00	4.360563×10^{-5}	1.00
$10^{-3}/16$	2.148817×10^{-5}	1.00	2.180284×10^{-5}	1.00
$10^{-3}/32$	1.074409×10^{-5}	1.00	1.090142×10^{-5}	1.00
$10^{-3}/64$	5.372044×10^{-6}	1.00	5.450714×10^{-6}	1.00

Table 8
Pressure convergence results for the linear penalty method (61) in Experiment 5, $r < 2$; grid parameter $h = 1/32$

ε	$\Vert p_h - p_h^{\varepsilon} \Vert_{0,r'}$ $(r = 3/2)$	Convergence rate	$\Vert p_h - p_h^{\varepsilon} \Vert_{0,r'}$ $(r = 7/4)$	Convergence rate
$10^{-3}/4$	7.703838×10^{-7}		1.047021×10^{-6}	
$10^{-3}/8$	3.851927×10^{-7}	1.00	5.235121×10^{-7}	1.00
$10^{-3}/16$	1.925965×10^{-7}	1.00	2.617564×10^{-7}	1.00
$10^{-3}/32$	9.629830×10^{-8}	1.00	1.308730×10^{-7}	1.00
$10^{-3}/64$	4.814916×10^{-8}	1.00	6.543918×10^{-8}	1.00

method for the power-law Stokes problem yields higher-order accuracy than the finite element discretization of the known nonlinear penalty method. The theoretical error estimates were supported by numerical experiments. We also proved the existence and uniqueness of solutions to the continuous and finite element discretization of the linear penalty method applied to the power-law Stokes problem. The extension of the improved linear penalty method to the Navier–Stokes equations with subgridscale artificial viscosity regularization will be the subject of a future study. The improvement in the linear penalty method could have a significant impact on turbulent flow computations, where computational efficiency is paramount.

Acknowledgements

The first author is partially supported by the National Science Foundation (DMS-0513542) and the Air Force Office for Scientific Research (F49620-03-1-0243 and FA9550-05-1-0449). The second author is partially supported by the National Science Foundation (DMS-0209309, DMS-0322852, and DMS-0513542) and the Air Force Office for Scientific Research (F49620-03-1-0243 and FA9550-05-1-0449). The third author is partially supported by the Air Force Office for Scientific Research (F49620-00-1-0299).

References

[1] R.A. Adams, Sobolev Spaces, Academic Press, New York, 1975.
[2] C. Amrouche, V. Girault, Decomposition of vector spaces and application to the Stokes problem in arbitrary dimension, Czech. Math. J. 44 (1994) 109–140.
[3] J. Baranger, P. Georget, K. Najib, Error estimates for a mixed finite element method for for a non-Newtonian flow, J. Non-Newtonian Fluid Mech. 23 (1987) 415–421.
[4] J. Baranger, K. Najib, Analyse numérique des écoulements quasi-Newtoniens dont la viscosité obéit à la loi puissance ou la loi de Carreau, Numer. Math. 58 (1990) 34–49.
[5] J.W. Barrett, W.B. Liu, Finite element error analysis of a quasi-Newtonian flow obeying the Carreau or power law, Numer. Math. 64 (4) (1993) 433–453.
[6] J.W. Barrett, W.B. Liu, Finite element approximation of the p-Laplacian, Math. Comp. 61 (1993) 523–537.
[7] L.C. Berselli, T. Iliescu, W.J. Layton, Mathematics of large eddy simulation of turbulent flows, in: Scientific Computation., Springer-Verlag, Berlin, 2006.
[8] D.V. Boger, A. Cabelli, A.L. Halmos, The behavior of a power-law fluid flowing through a sudden expansion, A.I.Ch.E.J. 21 (1975) 540–549.
[9] S.C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, Springer-Verlag, New York, 1994.
[10] S.S. Chow, Finite element error estimates for non-linear elliptic equations of monotone type, Numer. Math. 54 (1989) 373–393.
[11] C. Cuvelier, A. Segal, A.A. van Steenhoven, Finite Element Methods and Navier–Stokes Equations, D. Reidel Publishing Co., Boston, 1986.

[12] P. England, J. Jackson, Active deformation of the continents, Ann. Rev. Earth Planet. Sci. 17 (1989) 197–226.
[13] A. Ern, J.-L. Guermond, Theory and Practice of Finite Elements, in: Applied Mathematical Sciences, vol. 159, Springer-Verlag, New York, 2004.
[14] V. Girault, P.A. Raviart, Finite Element Approximation of the Navier–Stokes Equations, Springer-Verlag, New York, 1979.
[15] M.D. Gunzburger, Finite Element Methods for Viscous Incompressible Flows, Academic Press, Boston, 1989.
[16] T.J.R. Hughes, W.K. Liu, A. Brooks, Finite element analysis of incompressible viscous flows by the penalty function formulation, J. Comput. Phys. 30 (1) (1979) 1–60.
[17] L. Lefton, D. Wei, A penalty method for approximations of the stationary power-law Stokes problem, Electron. J. Differential Equations (7) (2001) 1–12.
[18] L. Lefton, D. Wei, Penalty finite element approximations of the stationary power-law Stokes problem, J. Numer. Math. 11 (2003) 301–322.
[19] J.L. Lions, Quelques Méthodes de Résolution des Problémes aux Limites Non Linéaires, Dunod, Paris, 1969.
[20] T.J. Liu, S. Wen, J. Tsou, Three-dimensional finite element analysis of polymeric fluid flow in an extrusion die. Part I: Entrance effects, Polymer Eng. Sci. 34 (10) (1994) 827–834.
[21] G. Minty, Monotone (nonlinear) operators in Hilbert space, Duke Math J. 29 (1961) 341–346.
[22] A. Quarteroni, A. Valli, Numerical Approximation of Partial Differential Equations, Springer-Verlag, Berlin, 1994.
[23] J.N. Reddy, V.A. Padhye, A penalty finite element model for axisymmetric flows of non-Newtonian fluids, Numer. Methods Partial Differential Equations 4 (1) (1988) 33–56.
[24] P. Sagaut, Large eddy simulation for incompressible flows, in: Scientific Computation, third edition, Springer-Verlag, Berlin, 2006.
[25] J. Shen, On error estimates of the penalty method for unsteady Navier–Stokes equations, SIAM J. Numer. Anal. 32 (2) (1995) 386–403.
[26] J.S. Smagorinsky, General circulation experiments with the primitive equations, Mon. Weather Rev. 91 (1963) 99–164.
[27] L.J. Sonder, P.C. England, Vertical averages of rheology of the continental lithosphere, Earth Planet. Sci. Lett. 77 (1989) 81–90.
[28] R. Temam, Une méthode d'approximation de la solution des équations de Navier–Stokes, Bull. Soc. Math. France 96 (1968) 115–152.
[29] D. Wei, Penalty approximations to the stationary power-law Navier–Stokes problem, Numer. Funct. Anal. Optim. 22 (2001) 749–765.